

Nucleotide Sequence of the *iucD* Gene of the pColV-K30 Aerobactin Operon and Topology of Its Product Studied with *phoA* and *lacZ* Gene Fusions

MARTA HERRERO, VICTOR DE LORENZO, AND J. B. NEILANDS*

Department of Biochemistry, University of California, Berkeley, California 94720

Received 13 July 1987/Accepted 29 September 1987

Gene *iucD* of the aerobactin operon of the *Escherichia coli* plasmid ColV-K30 encodes a membrane-bound enzyme synthesizing *N*⁶-hydroxylysine, the first product of the aerobactin biosynthesis pathway. The entire nucleotide sequence of the cloned *iucD* gene was determined, from which the primary and some aspects of the secondary structure of the encoded peptide were deduced. *E. coli* cells harboring multicopy plasmid pVLN12 (*iucD*⁺) hyperproduced an approximately 50-kilodalton peptide which was purified and identified as the product of the gene by examination of its amino-terminal sequence. Two *iucD*'-'*lacZ* gene fusions were constructed in vitro and four *iucD*'-'*phoA* gene fusions were generated in vivo by mutagenesis of *iucD* with transposon Tn*phoA* (Tn5 IS50_L::*phoA*). Analysis of the corresponding fusion proteins suggested at least two domains of attachment of the IucD protein to the inner side of the cytoplasmic membrane. The first apparent membrane-bound domain was found within the first 25 amino acids of the protein and showed a sequence which resembled that of the signal peptides.

Under conditions of iron stress, most bacteria induce high-affinity iron transport systems mediated by the chelators named siderophores (9, 42). The most thoroughly studied system is that encoded by the *Escherichia coli* virulence plasmid ColV-K30, which uses the siderophore aerobactin as an iron scavenger. This system is arranged in a single operon (11, 16, 24) driven by an iron-dependent promoter (6), which makes it an ideal model for study of the organization and transcriptional regulation of metal-sensing genes. Furthermore, the aerobactin system seems to be an important virulence determinant in nosocomial infections (for a review, see reference 14) and hence constitutes a potential target for new antimicrobial drugs.

The entire aerobactin system spans about 8 kilobases (kb) of pColV-K30, including a cluster of five genes transcribed coordinately through a strong promoter (6) located at the front of the first gene (Fig. 1). Transcription of the operon is controlled, as in other iron-regulated systems in *E. coli*, by the product of the *fur* gene (2, 26), a repressor protein which uses Fe(II) as corepressor to bind to an operator region in the aerobactin promoter (4, 17). Four of the five genes (*iucA* to *iucD*) are involved in the biosynthesis of the siderophore, while the remaining gene (*iutA*) determines an outer membrane receptor for the ferri-aerobactin complex. The function of each gene in subsequent steps of siderophore synthesis is known (16, 24), and their peptide products have been identified. In two cases (*iucB* and *iutA*), the proteins have been purified and their activity has been assessed in vitro (13, 34).

The product of the gene *iucD* is of special interest for two reasons. First, it catalyzes a unique type of enzymatic reaction which does not occur in animal tissues, namely, the *N*⁶-oxygenation of lysine. Second, *iucD* mediates the first step of aerobactin biosynthesis, and hence effective inhibitors of its activity may have therapeutic value.

A previous study (32) had shown the product of gene *iucD* to be bound to the cytoplasmic membrane, although the

function of the protein was uncertain at that time. In the present work we have determined the complete nucleotide sequence of the *iucD* gene and purified the corresponding peptide product. Structural features of the protein and details of its association to the inner membrane of *E. coli* were suggested by the analysis of *iucD*'-'*phoA* and *iucD*'-'*lacZ* gene fusions.

MATERIALS AND METHODS

Culture media, strains, and plasmids. Unless otherwise indicated, all cultures were grown at 37°C in LB medium (41) to which adequate levels of antibiotics (either ampicillin or ampicillin plus kanamycin) were added to ensure retention of the plasmids. Strains used are listed in Table 1. *E. coli* 71-18 was used as the host of all recombinant M13 clones for sequencing. *E. coli* LE392 was used for lytic growth of λ Tn*phoA* (Tn5 IS50_L::*phoA*), and *E. coli* CC118 was used as the recipient for Tn*phoA* mutagenesis of *iucD* (see below). Plasmid pVLN12 (Fig. 1) contains an approximately 3.1-kb *Bam*HI-*Eco*RI DNA fragment from the aerobactin operon (16) cloned at the corresponding sites of the pUC9 vector (52). This plasmid contains the complete sequence of *iucD* plus partial sequences of the flanking genes in the operon (17), *iucC*, and *iutA* (Fig. 1). The *iucD* gene in pVLN12 is expressed through the strong *lac* promoter of pUC9 (54). Plasmids resulting from the construction of *phoA* and *lacZ* gene fusions (see Table 2) will be described in the Results section.

DNA manipulation, enzymatic assays, and chemical determinations. All DNA techniques, including plasmid construction, ligation, preparation of competent *E. coli* cells, transformation, and fragment purification, were carried out by standard procedures (37) or following the directions of commercial suppliers. Restriction endonucleases were purchased from New England BioLabs, Inc., and "slow" *Bal* 31 nuclease (1 U/μl) was from International Biotechnologies, Inc. β-Galactosidase activity was measured by the method of Miller (41), and alkaline phosphatase activity was determined as previously described (40). The IucD⁺ phenotype

* Corresponding author.

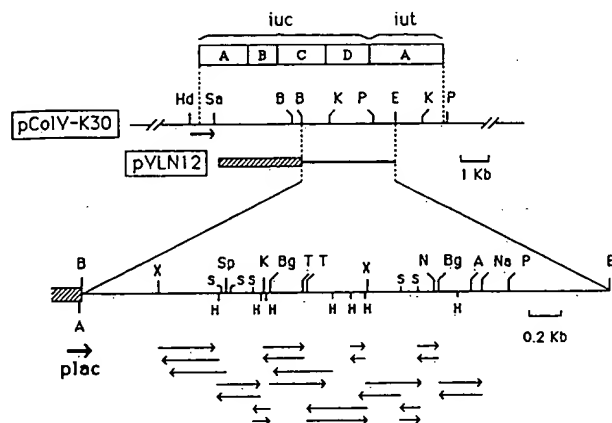


FIG. 1. Strategy for sequencing the *iucD* gene. Top: Genetic organization of the aerobactin operon lined up with a partial restriction map of the corresponding DNA segment of *pColV-K30*. The native iron-regulated promoter of the system (6) is located in front of *iucA* and is indicated by an arrow. Bottom: Enlarged representation of the portion of the aerobactin cluster cloned in *pVLN12* which contains the *iucD* gene. The orientation of the *lac* promoter from vector *pUC9* (through which the *iucD* gene is transcribed in *pVLN12*) is indicated, as are the restriction sites used to generate M13 clones for sequencing. Thin arrows indicate the location and orientation of the inserts in those M13 clones. Abbreviations: A, *AccI*; B, *BamHI*; Bg, *BglII*; E, *EcoRI*; Hd, *HindIII*; H, *HpaII*; K, *KpnI*; N, *NcoI*; Na, *NarI*; P, *PvuII*; S, *Sau3A1*; Sa, *Sall*; Sp, *SphI*; T, *TaqI*; X, *XmnI*.

was tested by assaying for hydroxylamine in the culture supernatants of the corresponding strains with the Csaky test, as modified by Gillam et al. (21).

Sequencing strategy and sequence analysis. The approximately 2.3-kb DNA fragment generated by treating *pVLN12* with the restriction endonuclease *AccI*, which overlaps most of the insert in *pVLN12* (Fig. 1), was purified and digested separately with the enzymes *Sau3A1* and *HpaII*. The resulting fragments were repurified from polyacrylamide gels (37) and cloned at compatible sites of the replicative forms of M13mp18 and M13mp19 phages. The inserts of single-stranded forms of the resulting constructions were sequenced by the dideoxynucleotide method (46). To confirm and complete the sequence of both DNA strands of the 2.3-kb *AccI* segment, *XmnI*, *SphI*, *KpnI*, *TaqI*, *BglII*, and *NarI* restriction endonucleases were used to generate additional fragments which were isolated, cloned, and sequenced by the same procedure (Fig. 1). The nucleotide sequence obtained was analyzed with a Sequence Analysis program from the Department of Biochemistry and Biophysics of the University of California at San Francisco and with a Mac Gene program from Applied Genetic Technology (Fairview Park, Ohio).

Protein techniques. Proteins hyperexpressed by *E. coli* strains carrying recombinant plasmids were isolated directly from whole-cell lysates of the corresponding strains as follows. After overnight growth in LB medium (41) with appropriate antibiotics, cells were spun down, suspended in sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) sample buffer (36), boiled for 3 min, and loaded in 8 to 10% polyacrylamide gels (36). The position of the protein bands of interest with respect to the migration of prestained molecular weight standards (Sigma Chemical Co.) was known in advance, and the corresponding slice of the unfixed gel was cut out and electroeluted in 80 mM TBE

buffer pH 8.3 (37) with an electroseparation device (Elutrap; Schleicher & Schuell) (30). The peptidic material recovered was precipitated in 10% trichloroacetic acid, washed in ethanol, and lyophilized. Analysis of the amino-terminal sequence of electroeluted protein bands was carried out in the Department of Biochemistry and Biophysics of the University of California at San Francisco.

Periplasmic proteins were released from *E. coli* cells by the chloroform shock procedure (18), dissolved in SDS-PAGE sample buffer (36), and analyzed in polyacrylamide gels as described above.

Spheroplasts from cells carrying gene fusion plasmids were prepared and analyzed as follows. Overnight cultures of the *E. coli* strains harboring the different plasmids were spun down and suspended in the same volume of a solution containing 2 mM EDTA, 0.58 M sucrose, and 30 mM Tris hydrochloride (pH 8). Lysozyme was then added to 100 μ g/ml. Conversion of cells into spheroplasts was monitored with a light microscope. Spheroplasts were spun down, suspended in an equal volume of the same buffer without lysozyme, and split into two parts, one of them being treated with trypsin (0.6 μ g/ml) (Sigma, T8642) for 30 min at 22°C. Treated and untreated spheroplasts were then spun down, suspended directly in SDS-PAGE sample buffer, boiled for 3 min, and analyzed in an SDS-PAGE system (36) as described above.

Construction of gene fusions. Alkaline phosphatase gene fusions to *iucD* were obtained in vivo by infecting *E. coli* CC118 harboring *pVLN12* (*iucD*⁺) with λ *TnphoA* as described (25). After infection, cells were plated on LB agar with kanamycin (300 μ g/ml) and 40 μ g of the chromogenic substrate 5-bromo-4-chloro-3-indolyl phosphate (XP; Sigma) per ml. High-kanamycin-concentration-resistant colonies were examined for the *PhoA*⁺ phenotype (blue in plates with XP) and for the production of hydroxylamine. Insertions of *TnphoA* into *pVLN12* were mapped by restriction enzyme analysis of the corresponding plasmids. When necessary, the precise boundaries of the *iucD*'-'*phoA* sequence junction were determined by sequencing, in an M13 system (see above), restriction fragments overlapping the junctions.

TABLE 1. Characteristics of relevant *E. coli* strains, phages and plasmids

Strain, phage, or plasmid	Genotype or description	Source or reference
<i>E. coli</i> 71-18	$\Delta(lac\ proAB)\ thi\ supE\ F'\ proAB\ lacI^qZ\ \Delta M15$	R. Kolter
CC118	$F^- \Delta(ara-leu)7697\ araD139\ \Delta lacX74\ galE\ galK\ phoA20\ thi\ rpsE\ rpoB\ argE(Am)recA1$	C. Manoil
LE392	$F^- supF\ supE\ hsdR\ galK\ trpR\ metB\ lacY\ tonA$	C. Manoil
λ <i>TnphoA</i>	b221 c1857 Pam3 with <i>TnphoA</i> in or near <i>rex</i>	C. Manoil
Plasmids		
pMLB1034	Ap' <i>lacZ</i> gene fusion vector	48
pVLN12	Ap' <i>iucD</i> ⁺	16
pOV7	Ap' <i>iucD</i> '-' <i>lacZ</i>	This study
pOV9	Ap' <i>iucD</i> '-' <i>lacZ</i>	This study
pOV5	Ap' Km' <i>iutA</i> '::TnphoA	This study
pOV25	Ap' Km' <i>iutA</i> '::TnphoA	This study
pOV33	Ap' Km' <i>iucD</i> '::TnphoA	This study
pOV99	Ap' Km' <i>iucD</i> '::TnphoA	This study
pOV48	Ap' Km' <i>iucD</i> '::TnphoA	This study
pOV28	Ap' Km' <i>iucD</i> '::TnphoA	This study

Two β -galactosidase gene fusions were obtained in vitro by ligating the 3.3-kb *Sma*I-*Bal*I fragment of pMLB1034 vector (48), containing a *lacZ* gene devoid of promoter and of ribosome-binding sequence, into plasmid pVLN12 digested with either *Kpn*I or *Nco*I and treated for a short time with *Bal* 31 enzyme. This last treatment was made to ensure the presence of all possible reading frames. The ligation mixtures were used to transform *E. coli* CC118. Transformant blue colonies arising in LB with 150 μ g of ampicillin and 40 μ g of 5-bromo-4-chloro-3-indolyl- β -D-galactoside (Sigma) solid medium per ml were picked for further analysis (see Results section).

RESULTS

Sequence of the *iucD* gene. pVLN12 (16) carries the active *iucD* gene in a 3.1-kb insert from pColV-K30 in pUC9. As the sequence of the following gene (*iutA*) is known (34) as well as the transcriptional orientation (6), we focused our attention on a 2.3-kb *Acc*I restriction fragment of pVLN12 in which the *iucD* gene sequence was expected to map. Figure 1 shows the collection of inserts in M13 clones which were used for sequencing the entire region. Each nucleotide was sequenced an average of three times.

Only one open reading frame of sufficient length to encode the IucD protein, previously sized in minicell systems in the range of 50 kilodaltons (kDa) (11, 16, 24, 32, 33), could be found with the transcriptional orientation of the aerobactin operon. The sequence of that open reading frame is shown in Fig. 2. This structural sequence (1,278 base pairs [bp]) contained an ATG start codon located 9 bp downstream from an AGGA sequence which might form part of a ribosome-binding site (47). The putative structural region spanned 426 amino acid residues and ended with the stop codon TAG. Of the residues, 48% were hydrophobic, 26% were polar uncharged, 15% were positively charged, and 11% were negatively charged. The predicted molecular weight of this peptide was 48,968. Figure 3 shows the hydrophobicity of the protein throughout the entire sequence according to Kyte and Doolittle (35).

The sequences upstream and downstream of the *iucD* structural gene were also examined. Most of the approximately 850 bp between the *Bam*HI site and the start of the *iucD* gene overlapped the structural sequence of the preceding gene in the operon, *iucC*, although the precise boundaries are unknown. No evident promoterlike sequences, i.e., -35 and -10 consensus sequences, were detected either up- or downstream of *iucD*. Interestingly, the intercistronic regions at both sides of *iucD* contained short regions of dyad symmetry (Fig. 2). The dyad preceding the *iucD* structural gene would allow the formation of a stem-loop structure in the corresponding transcript with a calculated ΔG value of -20.6 kcal/mol, while the dyad succeeding the gene would have a ΔG of -17.6 kcal/mol.

Isolation and identification of the *iucD* protein product. *E. coli* cells harboring pVLN12 hyperproduced an approximately 50-kDa peptide when analyzed in SDS-PAGE systems (not shown). This polypeptide had a molecular weight similar to that of a protein identified in minicells as the product of the *iucD* gene (16, 24). To ascertain whether this peptide was indeed the product of the gene and to check the reading frame obtained from nucleotide sequence data, we purified this protein band and subjected it to amino-terminal analysis. *E. coli* CC118 harboring the plasmid pOV25 was used as the source of the protein (see below). For unknown reasons, this plasmid construction afforded further hyperex-

pression of the approximately 50-kDa protein to a level close to 5% of the total cell protein (see Fig. 5). The protein band was purified from a whole-cell extract by electroelution from a polyacrylamide gel. The initial 15 amino acids of the isolated polypeptide were identified as NH₂-Met-Lys-Lys-Ser-Val-Asp-Phe-Ile-Gly-Val-Gly-Thr-Gly-Pro-Phe. These are the same residues expected from the translated sequence, including the initial methionine as the amino terminus. This result allowed the unequivocal assignment of the approximately 50-kDa protein as the product of the *iucD* gene.

***phoA* and *lacZ* gene fusions to *iucD*.** Alkaline phosphatase in *E. coli* must be exported into the periplasm to be active (28, 40). Therefore, *phoA* gene fusions are a convenient method to probe the topology of putative membrane-bound and exported proteins (7, 38, 39). To construct a number of *iucD*'-*phoA* gene fusions, we mutagenized an *E. coli* strain harboring the plasmid pVLN12 with *TnphoA* and selected insertions of the transposon in the multicopy plasmid with a high concentration of kanamycin. Should the product of *iucD* be a membrane-associated protein and should it have domains facing the periplasm, colonies with a *PhoA*⁺ phenotype concomitant with the loss of hydroxylamine production (*iucD*) should be isolated after mutagenesis. A number of blue colonies in XP medium indeed appeared after *TnphoA* mutagenesis, and they were picked for further analysis. Most of the deep-blue colonies resulted from insertions in the sequence of *iutA* present in pVLN12. This was expected since the truncated sequence of *iutA* remaining in pVLN12 (Fig. 4) keeps the intact signal peptide sequence necessary for protein export (34). The appearance of these conspicuous *PhoA*⁺ colonies due to insertions in *iutA* constituted an internal control for the performance of the procedure.

Figure 4 summarizes the location and orientation of the *TnphoA* insertions found. Four insertion derivative plasmids, pOV33, pOV99, pOV48, and pOV28, conferred a *PhoA*⁺ *iucD*⁻ phenotype and were studied in more detail. As a complement to the *phoA* fusions, we also obtained two *iucD*'-*lacZ* gene fusions in vitro as described in the Materials and Methods section, which were termed pOV7 and pOV9. The precise positions of fusion junctions were confirmed by nucleotide sequencing, and they are indicated in Fig. 3 and Table 2.

Analysis of IucD'-PhoA and IucD'-LacZ fusion proteins. The levels of alkaline phosphatase and β -galactosidase afforded by *E. coli* cells carrying different *phoA* and *lacZ* fusions are shown in Table 2. All four *iucD*'-*phoA* fusions (carried in pOV33, pOV99, pOV48, and pOV28) afforded phosphatase activity, although the levels varied significantly in each fusion. Cells harboring pOV33 and pOV99 showed 15- to 20-fold-higher levels of phosphatase than those with pOV48 and pOV28, while an *iutA*'-*phoA* fusion used as a positive control of effective export (pOV5) had levels fivefold above those of the most active *iucD*'-*phoA* fusion (pOV33).

Plasmids with *iucD*'-*lacZ* fusions (pOV7 and pOV9) afforded rather high values for β -galactosidase activity (Table 2). Permeabilization of cells prior to the assay resulted in a 16-fold increase in β -galactosidase activity detected in cells carrying pOV7, while the same treatment produced an increase of only 2.5-fold in those with pOV9.

Figure 5 shows the protein profile of *E. coli* cells harboring the plasmids with the *iucD*'-*phoA* and *iucD*'-*lacZ* fusions analyzed in an SDS-PAGE system. In all cases, hyperproduction of a protein of the approximate size of the predicted

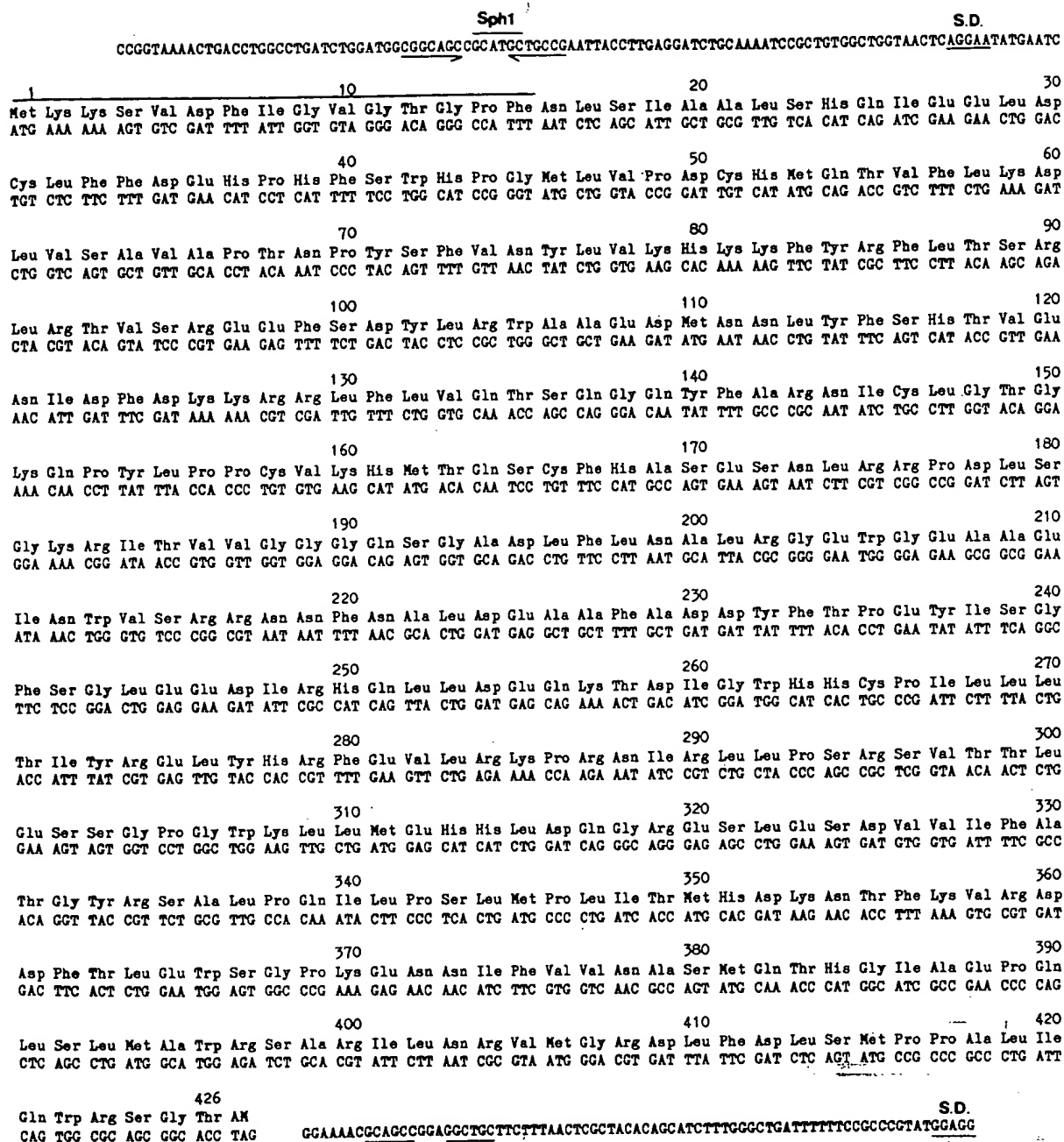


FIG. 2. Nucleotide sequence of *iucD*. Nucleotide sequence of the DNA region of pVLN12 containing the *iucD* gene of the aerobactin operon. The amino acid sequence of the predicted translation product is shown above the nucleotide sequence. The numbers above each line refer to the number of amino acid residues. The segment marked with a line corresponds to the amino-terminal residues of the purified *iucD* product determined by protein sequencing (see text). Two potential Shine-Dalgarno (S.D.) ribosome-binding sites (47) are indicated. The one assigned to *iucD* is located 9 bp upstream of the initiation codon ATG; the signal at the end of the sequence has been previously ascribed to *iutA*, the next gene in the aerobactin operon (34). Two short regions of dyad symmetry located in the intercistronic sequences surrounding the *iucD* gene are shown by paired arrows. The unique *SphI* restriction site in pVLN12 is also indicated.

fusion product (Table 2) was observed. Incidentally, we observed that cells containing a fusion plasmid named pOV25 excreted increased amounts of hydroxylamine into the culture (as detected with the Csaky test modified by Gillam et al. [21]) and produced a stronger 50-kDa band when analyzed in an SDS-PAGE system. pOV25 consisted of a pVLN12 derivative with a *TnpH* insertion in *iutA*

oriented opposite to *iucD* transcription (Fig. 4). *E. coli* cells carrying this construction were subsequently used as the source of the approximately 50-kDa protein.

To check whether the *IucD*'-PhoA and *IucD*'-LacZ fusion proteins could be entirely located in the periplasm, the corresponding *E. coli* cells were subjected to chloroform shock to release periplasmic proteins (18). No fusion protein

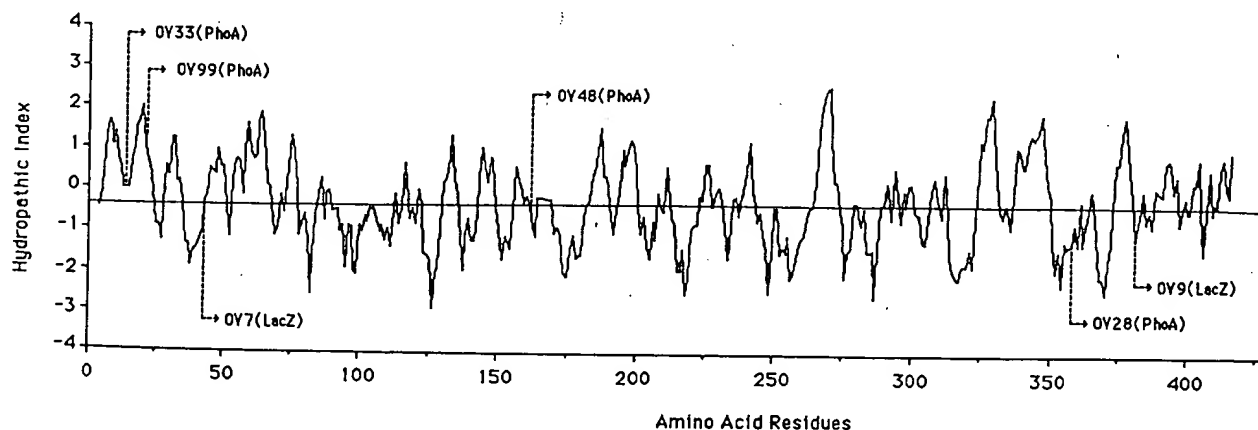


FIG. 3. Hydropathicity profile of *lucD* protein. The hydropathic index was determined at a span setting of 7 residues with the algorithm presented by Kyte and Doolittle (35). Hydrophobic regions extend above the middle line. The arrows show the location of the inserts leading to the four *iucD'*-*phoA* and the two *iucD'*-*lacZ* gene fusions studied in this work.

was observed in the shock fluid (not shown), ruling out a free periplasmic location.

To ascertain whether domains of the fusion proteins could be accessible from the outside of the cytoplasmic membrane, spheroplasts from every fusion-containing *E. coli* strain (Table 2) were prepared, treated with a low concentration of trypsin, and subsequently analyzed in SDS-PAGE systems. The protein fusion bands from pOV33, pOV99, pOV48, pOV28, and pOV9 faded from spheroplast proteins after trypsin treatment, while that of pOV7 remained insensitive to protease (Fig. 6). Interestingly, in spheroplasts from *E. coli* carrying the intact *iucD* gene, the approximately 50-kDa product was not significantly affected by trypsin (Fig. 6).

DISCUSSION

Results from different groups (16, 24) support the current hypothesis that the biosynthetic route of the siderophore aerobactin in pColV-enclosed systems is identical to that proposed by Appanna et al. (1) for *Aerobacter aerogenes* 62-I, the microorganism for which the production of aerobactin was first described. In this scheme, the very first step leading to aerobactin production is the *N*⁶-hydroxylation of L-lysine, followed by an *N*⁶-acylation of the corresponding *N*⁶-hydroxylysine and an eventual condensation of two *N*⁶-hydroxy-*N*⁶-acetyl lysine moieties with a citric acid backbone (1, 16, 24).

The presence of an approximately 50-kDa membrane-bound protein as the product of one of the genes of the aerobactin system of pColV-K30 was first shown by Krone et al. (32) and assigned a role in the transport of aerobactin. However, overwhelming evidence from different laboratories later demonstrated that the peptide was instead involved in aerobactin biosynthesis (11, 44), in particular in the step of *N*⁶-hydroxylation of L-lysine (16, 24) which precedes the eventual synthesis of the aerobactin molecule. This assignment, based on complementation analysis (23) and studies on accumulation of aerobactin precursors by different mutants (16, 24), is consistent with recent *in vitro* studies made with the unpurified protein (27). In addition, Coy et al. (13) showed that *N*⁶-hydroxylysine is the substrate for the purified acetylase, the enzyme which mediates the next step in the biosynthetic pathway. Some remaining uncertainties (20) in the biosynthetic pathway of aerobactin synthesis have been addressed elsewhere (3). The gene of the aerobactin

operon encoding this protein was termed *iucD* (16) or, alternatively, *aerA* (23).

The complete sequence of the gene is displayed in Fig. 2 and to our knowledge is the first reported for a biosynthetic gene of a siderophore. The sequence accounts for the production of a predicted 48,968-dalton peptide, a value very close to the approximately 50 kDa observed in mini- and maxicell systems by several laboratories (11, 16, 24, 32, 33).

IucD protein is a moderately hydrophobic polypeptide. Analysis of the 15 amino-terminal residues of the gene product reveals that mature *IucD* protein is the unprocessed translational product of the *iucD* gene sequence, including the initial ATG codon. The amino-terminal sequence of *IucD* also resembles that of known signal peptides in that it starts with basic amino acids (lysines at positions 2 and 3) followed by a region rich in hydrophobic residues (Fig. 2). Uncleaved leaderlike sequences are frequently found in inner membrane proteins of *E. coli*, as for instance in the Tet protein from Tn10 or pBR322 (43), the Tsr protein (serine chemoreceptor [8]), the Tar protein (aspartate chemoreceptor [45]), and others.

Analysis of codon usage in *iucD* suggests that it is a rather weakly expressed gene since nonoptimal codons are frequent throughout the sequence, thus possibly reducing the rate of translation. A remarkable feature is the appearance at positions corresponding to amino acids 11, 12, and 13 of a group of three codons (GGG, ACA, and GGG) which are

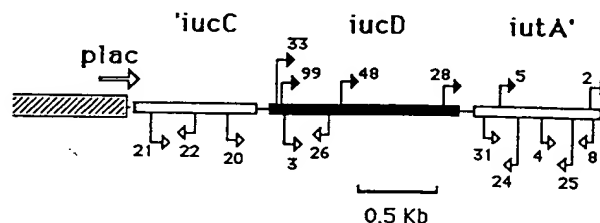


FIG. 4. Positions of *TnphoA* insertions in plasmid pVLN12. The open arrows indicate the position and orientation of different *TnphoA* insertions throughout the 3.1-kb insert of pVLN12, the derivative plasmids being identified with the same pOV serial number (see Tables 1 and 2). Solid arrowheads represent insertions that generate in-frame translational sequences. The striped segment of the figure symbolizes the vector portion of constructions.

TABLE 2. Characteristics of 'phoA' and 'lacZ' gene fusions in *E. coli* CC118

Plasmid ^a	Gene fusion	AP activity ^b (U/OD ₆₀₀)	No. of IucD residues in hybrid ^c	Protein size (kDa)		Hydroxylamine production ^f	β-Galactosidase activity ^g (U/OD ₆₀₀)	
				Predicted ^d	Found ^e		Permeabilized cells	Intact cells
pOV25	None	≤1	426	49	47.5	+		
pOV33	<i>iucD'</i> - <i>'phoA</i>	375	14	48.5	50	-		
pOV99	<i>iucD'</i> - <i>'phoA</i>	110	21	49.5	51	-		
pOV48	<i>iucD'</i> - <i>'phoA</i>	12	160	65	61	-		
pOV28	<i>iucD'</i> - <i>'phoA</i>	10	357	88	87	-		
pOV5	<i>iutA'</i> - <i>'phoA</i>	1,700	N.A.	N.D.	57	+		
pOV7	<i>iucD'</i> - <i>'lacZ</i>		42	120	120	-	15,000	1,100
pOV9	<i>iucD'</i> - <i>'lacZ</i>		380	159	152	-	1,500	580

^a Construction and structure of plasmids pOV25, pOV33, pOV99, pOV48, pOV28, pOV5, pOV7, and pOV9 are described in the text and in Table 1. Plasmids were harbored in all cases by *E. coli* CC118.

^b Alkaline phosphatase (AP) activity was assayed by measuring the rate of *p*-nitrophenylphosphate hydrolysis in permeabilized cells (40). *E. coli* CC118 cells carrying the different plasmids were grown in LB liquid culture, washed, and suspended in medium 121 (31) prior to the assay. Values are the averages of three independent experiments.

^c Number of amino acid residues of the IucD product present in the hybrid protein. Every DNA junction fragment was cloned in phage M13 derivatives; its nucleotide sequence was obtained by the dideoxynucleotide chain termination method (46), and from this the amino acid sequence was deduced. N.A., Not applicable (pOV5 contains an *iutA'*-*'phoA* gene fusion and a complete *iucD* gene).

^d Calculated from the position of the *TnphoA* insertion, assuming 47 kDa as the molecular weight of alkaline phosphatase in *TnphoA* (38), and an average molecular weight of 115 for the amino acid residue in the IucD product. In the case of *iucD'*-*'lacZ* fusions, 116 kDa was the mass considered the β-galactosidase contribution (41). N.D., Not determined.

^e Determined for hybrid proteins separated by SDS-PAGE by their rates of migration relative to proteins of known molecular weight.

^f Detection of hydroxylamine-related compound(s) in culture supernatants by the Csaky test as modified by Gillam et al. (21).

^g *E. coli* CC118(pOV7) and *E. coli* CC118(pOV9) cells were washed in Z buffer (41), and their β-galactosidase activity was determined either on cells permeabilized with chloroform and SDS (total activity) or on intact cells to estimate the activity independent of Lac Y permeation (49).

very unusual in the normal codon usage of *E. coli* (22). Clustering of rare codons in signal sequences seems to further decrease the rate of translation (51), which is believed to facilitate early interactions of proteins with the inner membrane (10).

Some features of the mode of association of this protein with the inner membrane were suggested by the study of a number of *iucD'*-*'phoA* and *iucD'*-*'lacZ* gene fusions. The rationale for this approach has been discussed in detail elsewhere (38-40). In summary, alkaline phosphatase is inactive when localized in the cytoplasm (28, 40), and therefore, the activity of a *'phoA* gene fusion will depend on the translocation of the hybrid protein to the periplasm, or at least the *'PhoA* moiety of the fusion must face the periplasm (38, 39). A similar but less stringent approach can be applied to *'lacZ* gene fusions carried by *lacY* mutants (defective in lactose permease). In these cells, a hybrid protein localized in the cytoplasmic membrane will afford a remarkable increase in β-galactosidase activity (determined by hydrolysis of *o*-nitrophenyl-β-D-galactoside) after permeabilization of the cells if the *'LacZ* moiety of the protein is cytoplasmic, while a smaller difference will be observed if it faces the periplasm (50).

Four *iucD'*-*'phoA* fusions conferring a *PhoA*⁺ phenotype were isolated (Table 2). All of them were stable and constituted a major protein component of the cells carrying the corresponding plasmids (Fig. 6). This alone rules out the free cytoplasmic location of the fusions, since alkaline phosphatase is generally quite unstable when exposed to the cytoplasm (38). The hybrid proteins encoded by pOV33 (at the 14th amino acid residue of IucD) and pOV99 (at residue 21) (Fig. 3 and Table 2) afforded a significant level of alkaline phosphatase activity. These proteins did not appear in the periplasmic fluid after chloroform shock, suggesting that the *'PhoA* moiety of the fusion is anchored in the inner membrane and facing the periplasm. Since no information required to act as an export signal or to block export is found in either the alkaline phosphatase or the 17 residues contributed by the IS50_L of *TnphoA* (38), we conclude that the

translocation to the periplasm of *'PhoA* moieties and the final anchoring at the membrane of the two smaller IucD'-*'PhoA* proteins is promoted by the sequence of the first 14 residues from IucD present in both fusions. It is likely that part of these 14 residues forms a thermodynamically stable association with the membrane, probably following a model of membrane-spanning domain closely related to those discussed by Davis et al. (15). These authors have shown that not only hydrophobic α-helices but also strings containing only 11 hydrophobic amino acids can generate membrane-spanning domains. Recent studies have also demonstrated that an anchoring domain can be generated by a short stretch of hydrophobic and neutral residues (19), and even the five to eight consecutive hydrophobic amino acids near the N

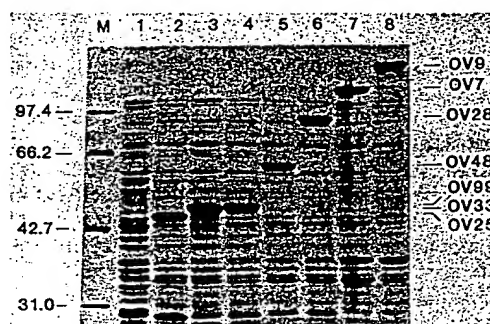


FIG. 5. Analysis of hybrid proteins synthesized in *E. coli* CC118 cells carrying different fusion plasmids. The proteins present in crude extracts of cells were separated by SDS-PAGE (8% acrylamide) and stained with Coomassie blue. Lanes: M, molecular weight markers (in thousands); 1, control *E. coli* CC118, no plasmid; 2, control *E. coli* CC118(pOV25) (pOV25 carries a *TnphoA* insertion in *iutA* which leaves intact *iucD*; see Fig. 4); 3, *E. coli* CC118(pOV33); 4, *E. coli* CC118(pOV99); 5, *E. coli* CC118(pOV48); 6, *E. coli* CC118(pOV28); 7, *E. coli* CC118(pOV7); 8, *E. coli* CC118(pOV9). The protein products IucD, the four IucD'-*'PhoA* fusions, and the two IucD'-*'LacZ* fusions can be identified as the major bands in lanes 2, 3 to 6, and 7 to 8, respectively.

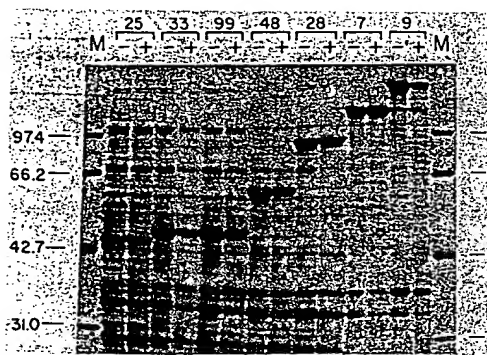


FIG. 6. Accessibility of *IucD*'-PhoA and *IucD*'-LacZ protein fusions to trypsin in spheroplasts. A Coomassie-blue stained SDS-PAGE system (8% acrylamide) loaded with spheroplasts from cells harboring gene fusion plasmids is shown. Numbers on top refer to the number of the OV fusion plasmid carried in each case by *E. coli* CC118 cells. Treatment of the spheroplasts with trypsin is indicated above each lane. Cells with the construction pOV25 were introduced as a control: pOV25 carries a *TnphoA* insertion in *iutA* which leaves intact *iucD* (see Fig. 5). Lanes M contained molecular weight standards (in thousands).

terminus of a 'LacZ protein fusion can serve to anchor a functional β -galactosidase to the inner membrane of *E. coli* (12).

Since the leading region of 14 and 21 amino acids in the fusions OV33 and OV99, respectively, appears to be enough to promote the export of their PhoA moiety, it is likely that the leading peptide is also able to promote interactions of the *IucD* product itself with the membrane. The leading region of at least 14 amino acids might define a first domain of attachment of this protein to the cytoplasmic membrane.

The levels of alkaline phosphatase afforded by the PhoA⁺ fusions encoded by pOV48 (*TnphoA* insertion at residue 160 of *IucD*) and pOV28 (at residue 357) were much lower than the levels detected in the presence of pOV33 and pOV99 (Table 2), although the corresponding proteins were equally stable and synthesized in similar quantities (Fig. 5).

Two *IucD*'-lacZ gene fusions were prepared in vitro by ligating the 'lacZ gene to residues 42 and 380 of *IucD* sequence in plasmid pVLN12, generating, respectively, plasmids pOV7 and pOV9. Both of them yielded rather high β -galactosidase activities in *E. coli* CC118 cells, which lack lactose permease. It is remarkable that permeabilized cells with fusion plasmid pOV9 increased only 2.5 times their β -galactosidase activity levels compared with intact cells, while the corresponding increase detected in cells with pOV7 was over 12-fold. In *lacY* cells the activity of cytoplasmic free β -galactosidase increased usually over 10 times after permeabilization (not shown). β -Galactosidase values suggested that the 'LacZ moiety in the pOV9 protein fusion is facing the periplasm, while that of pOV7 is facing the cytoplasm. Both hybrid proteins were stable and synthesized in similar amounts, but the total β -galactosidase activity levels were some 10-fold higher with pOV7 than with pOV9.

A final set of assays was carried out to check whether domains of the *IucD*'-PhoA and *IucD*'-LacZ fusion proteins could be facing the periplasm. For this purpose, spheroplasts from cells carrying the fusions were prepared and treated with trypsin, and their protein profiles were compared with those of untreated spheroplasts (Fig. 6). Although spheroplasts from cells carrying pOV25 with the functional *iucD* gene did not reveal a significant degradation of the approximately 50-kDa *IucD* protein, all four *IucD*'-

PhoA fusions appeared to be sensitive to trypsin, since analysis in an SDS-PAGE system showed that the corresponding bands disappeared at high trypsin concentrations (not shown) or at least faded after protease treatment (low trypsin concentration, Fig. 6). With respect to the *IucD*'-LacZ fusions, Fig. 6 also shows that the fusion from pOV9 was heavily degraded, while that of pOV7 remained virtually intact. These last two fusions (OV7 and OV9) provided reasonable controls for the performance of the procedure, because cytoplasmic PhoA fusions are unstable (38) and would therefore be unreliable as controls in this type of experiment.

It is known that fusions of some exported proteins to 'LacZ are lethal to *E. coli* cells when produced in large amounts, a phenomenon explained as jamming of the export machinery and concomitant accumulation of precursor forms of other exported proteins (5, 29, 40, 50). Although the signal sequences of the corresponding proteins (i.e., alkaline phosphatase, maltose-binding protein, etc.) are capable of directing β -galactosidase to the membrane, the export is not completed and the hybrid proteins become membrane bound. Since the 'LacZ moiety was never found translocated into the periplasm, it was considered that β -galactosidase is not competent to cross through the lipid bilayer (40). In our case, the small difference found in β -galactosidase activity between permeabilized and intact cells containing plasmid pOV9, together with the apparent accessibility of the same pOV9-encoded hybrid protein to trypsin digestion in spheroplasts, indicate that the 'LacZ moiety has been mobilized into the periplasm, and therefore that β -galactosidase can indeed be translocated. This also shows that membrane-bound proteins fused to 'LacZ can be synthesized in large amounts, apparently without deleterious consequences. Probably the *IucD* product is localized in or at the inner membrane independently of the export machinery, and this fact circumvents the problem arising from blockage of the export pathway detected in the study of some other hybrid proteins (5, 29, 40).

The previous results support the association of the *iucD* gene product with the inner membrane of *E. coli* cells, with the bulk of the protein facing the inner side and with at least two attachment sites which extend into the lipid bilayer. One of them would be at the leading sequence of the protein and is defined by the 'phoA fusions of pOV33 and pOV99, while the other occurs towards its terminal portion and is defined by the lacZ fusion of pOV9. This last domain may also account for the 'phoA fusion of pOV28. Although the peptide sequences which probably protrude into the membrane afford translocation of the 'PhoA and 'LacZ moieties of those fusions into the periplasm, there is no evidence that the *IucD* product itself has any domain exposed to the periplasm. An additional region of attachment to the membrane could occur in the region close to residue 160, accounting for the PhoA fusion of pOV48, since cells with that plasmid show detectable alkaline phosphatase activity and the corresponding fusion protein is sensitive to trypsin in spheroplasts. The low activity of the *IucD*'-PhoA proteins encoded by pOV48 and pOV28 may result from an unusual type of association with the membrane that could affect some of the steric requirements for activation of alkaline phosphatase (39).

The location of the *IucD* product makes biological sense. As this enzyme is involved in iron metabolism and probably requires molecular oxygen for its activity (3), its membrane location would contribute to maintenance of a reducing intracellular milieu.

Important features of the gene product of *iucD* remain uncertain, mostly those regarding its enzymology. We think that the information on the gene sequence and the characteristics of the protein described here will contribute to the elucidation of a rational scheme for its purification and, ultimately, to the discovery of inhibitors of its activity.

ACKNOWLEDGMENTS

This research was supported in part by Public Health Service grant A104156 from the National Institutes of Health (NIH) and by grants PCM78-12198 and CRCR-1-1633 from the National Science Foundation and Department of Agriculture, respectively. M.H. was a Postdoctoral Fellow from the Fulbright-MEC/Spain Program. V. de L. was an International Fellow of the Fogarty Center of the National Institutes of Health (1F05TW03577-01-BI-5).

This work was made possible thanks to the reliable advice of Roberto Kolter on sequencing procedures and the valuable strains prepared by Colin Manoil and kindly provided by Kathy Strauch, both from Jon Beckwith's laboratory. We are also indebted to Fabio Giovannini and Christian Leumann for their essential chemical advice.

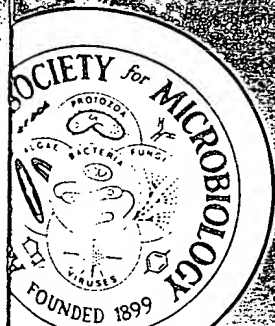
LITERATURE CITED

- Appanna, D. I., B. J. Grundy, E. W. Szczepan, and T. Viswanatha. 1984. Aerobactin synthesis in a cell-free system of *Aerobacter aerogenes* 62-1. *Biochim. Biophys. Acta* 801:437-443.
- Bagg, A., and J. B. Neilands. 1985. Mapping of a mutation affecting regulation of iron uptake systems in *Escherichia coli* K-12. *J. Bacteriol.* 161:450-453.
- Bagg, A., and J. B. Neilands. 1987. Molecular mechanism of regulation of siderophore-mediated iron assimilation. *Microbiol. Rev.* 51:509-518.
- Bagg, A., and J. B. Neilands. 1987. The Fur protein acts as repressor employing Fe(II) as a cofactor to bind the operator of an iron transport operon in *Escherichia coli*. *Biochemistry* 26:5471-5477.
- Bassford, P. J., T. J. Silhavy, and J. R. Beckwith. 1979. Use of gene fusion to study secretion of maltose-binding protein into *Escherichia coli* periplasm. *J. Bacteriol.* 139:19-31.
- Bindereif, A., and J. B. Neilands. 1985. Promoter mapping and transcriptional regulation of the iron assimilation system of plasmid ColV-K30 in *Escherichia coli* K-12. *J. Bacteriol.* 162:1039-1046.
- Boquet, P. L., C. Manoil, and J. Beckwith. 1987. Use of *TnphoA* to detect genes for exported proteins in *Escherichia coli*: identification of the plasmid-encoded gene for a periplasmic acid phosphatase. *J. Bacteriol.* 169:1663-1669.
- Boyd, A., K. Kendall, and M. I. Simon. 1983. Structure of the serine chemoreceptor in *Escherichia coli*. *Nature (London)* 301:623-626.
- Braun, V. 1985. The unusual features of the iron transport systems of *Escherichia coli*. *Trends Biochem. Sci.* 10:75-78.
- Burns, D. M., and I. R. Beacham. 1985. Rare codons in *E. coli* and *S. typhimurium* signal sequences. *FEBS Lett.* 189:318-324.
- Carbonetti, N. H., and P. H. Williams. 1984. A cluster of five genes specifying the aerobactin iron uptake system of plasmid ColV-K30. *Infect. Immun.* 46:7-12.
- Conway, T., Y. A. Osman, and L. O. Ingram. 1987. Gene expression in *Zymomonas mobilis*: promoter structure and identification of membrane anchor sequences forming functional LacZ' fusion proteins. *J. Bacteriol.* 169:2327-2335.
- Coy, M., B. H. Paw, A. Bindereif, and J. B. Neilands. 1986. Isolation and properties of N^ε-hydroxylysine:acetyl coenzyme A N^ε-transacetylase from *Escherichia coli* pABN11. *Biochemistry* 25:2485-2489.
- Crosa, J. H. 1984. The relationship of plasmid-mediated iron transport and bacterial virulence. *Annu. Rev. Microbiol.* 38:69-89.
- Davis, N. G., J. D. Boeke, and P. Model. 1985. Fine structure of a membrane anchor domain. *J. Mol. Biol.* 181:111-121.
- De Lorenzo, V., A. Bindereif, B. H. Paw, and J. B. Neilands. 1986. Aerobactin biosynthesis and transport genes of plasmid ColV-K30 in *Escherichia coli* K-12. *J. Bacteriol.* 165:570-578.
- De Lorenzo, V., S. Wee, M. Herrero, and J. B. Neilands. 1987. Operator sequences of the aerobactin operon of plasmid ColV-K30 binding the ferric uptake regulation (*fur*) repressor. *J. Bacteriol.* 169:2624-2630.
- Ferro-Luzzi Ames, G., C. Prody, and S. Kustu. 1984. Simple, rapid, and quantitative release of periplasmic proteins by chloroform. *J. Bacteriol.* 160:1181-1183.
- Fikes, J. D., and P. J. Bassford. 1987. Export of unprocessed precursor maltose-binding protein to the periplasm of *Escherichia coli* cells. *J. Bacteriol.* 169:2352-2359.
- Ford, S., R. A. Cooper, and P. H. Williams. 1986. Biochemical genetics of aerobactin biosynthesis in *Escherichia coli*. *FEMS Microbiol. Lett.* 36:281-285.
- Gillam, A. H., A. G. Lewis, and R. J. Andersen. 1981. Quantitative determination of hydroxamic acids. *Anal. Chem.* 53:841-844.
- Grosjean, H., and W. Fiers. 1982. Preferential codon usage in prokaryotic genes: the optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes. *Gene* 18:199-209.
- Gross, R., F. Engelbrecht, and V. Braun. 1984. Genetic and biochemical characterization of the aerobactin synthesis operon on pColV. *Mol. Gen. Genet.* 196:74-80.
- Gross, R., F. Engelbrecht, and V. Braun. 1985. Identification of the genes and their polypeptide products responsible for aerobactin synthesis by pColV plasmids. *Mol. Gen. Genet.* 201:204-212.
- Gutierrez, C., J. Barondess, C. Manoil, and J. Beckwith. 1987. The use of transposon *TnphoA* to detect genes for cell envelope proteins subject to a common regulatory stimulus. Analysis of osmotically regulated genes in *Escherichia coli*. *J. Mol. Biol.* 195:289-297.
- Hantke, K. 1981. Regulation of ferric iron transport in *Escherichia coli* K-12: isolation of a constitutive mutant. *Mol. Gen. Genet.* 182:288-292.
- Heydel, P., H. Plattner, and H. Dieckmann. 1987. Lysine N-hydroxylase and N-acetyltransferase of the aerobactin system of pColV plasmids in *Escherichia coli*. *FEMS Microbiol. Lett.* 40:305-309.
- Hoffman, C. S., and A. Wright. 1985. Fusions of secreted proteins to alkaline phosphatase: an approach for studying protein secretion. *Proc. Natl. Acad. Sci. USA* 82:5107-5111.
- Ito, K., P. J. Bassford, Jr., and J. Beckwith. 1981. Protein localization in *E. coli*: is there a common step in the secretion of periplasmic and outer-membrane proteins? *Cell* 24:707-717.
- Jacobs, E., and A. Clad. 1986. Electroelution of fixed and stained membrane proteins from preparative sodium dodecyl sulfate-polyacrylamide gels into a membrane trap. *Anal. Biochem.* 154:583-589.
- Kreuzer, K., C. Pratt, and A. Torriani. 1975. Genetic analysis of regulatory mutants of alkaline phosphatase of *E. coli*. *Genetics* 81:459-468.
- Krone, W. J. A., J. Luirink, G. Koningstein, B. Oudega, and F. K. de Graaf. 1983. Subcloning of the cloacin DF-13/aerobactin receptor protein and identification of a pColV-K30-determined polypeptide involved in ferric aerobactin uptake. *J. Bacteriol.* 156:945-948.
- Krone, W. J. A., B. Oudega, F. Stegehuis, and F. K. de Graaf. 1983. Cloning and expression of the cloacin DF-13/aerobactin receptor of *Escherichia coli* (pColV-K30). *J. Bacteriol.* 153:716-721.
- Krone, W. J. A., F. Stegehuis, G. Koningstein, C. van Doorn, B. Roosendaal, F. K. de Graaf, and B. Oudega. 1985. Characterization of the pColV-K30 encoded cloacin DF/13 aerobactin outer membrane receptor protein of *Escherichia coli*: isolation and purification of the protein and analysis of its nucleotide sequence and primary structure. *FEMS Microbiol. Lett.* 26:153-161.
- Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.*

- 157:105-132.
36. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* 227:680-685.
37. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
38. Manoil, C., and J. Beckwith. 1985. *TnphoA*: a transposon probe for protein export signals. *Proc. Natl. Acad. Sci. USA* 82:8129-8133.
39. Manoil, C., and J. Beckwith. 1986. A genetic approach to analyzing membrane protein topology. *Science* 233:1403-1408.
40. Michaelis, S., H. Inouye, D. Oliver, and J. Beckwith. 1983. Mutations that alter the signal sequence of alkaline phosphatase in *Escherichia coli*. *J. Bacteriol.* 154:366-374.
41. Miller, J. H. 1972. *Experiments in molecular genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
42. Neilands, J. B. 1981. Iron absorption and transport in microorganisms. *Annu. Rev. Nutr.* 1:27-46.
43. Nguyen, T. T., K. Postle, and P. Bertrand. 1983. Sequence homology between the tetracycline-resistance determinants of *Tn10* and *pBR322*. *Gene* 25:83-92.
44. Roberts, M., R. W. Leavitt, N. H. Carbonetti, S. Ford, R. A. Cooper, and P. H. Williams. 1986. RNA-DNA hybridization analysis of transcription of the plasmid ColV-K30 aerobactin gene cluster. *J. Bacteriol.* 167:467-472.
45. Russo, A. F., and D. E. Koshland. 1983. Separation of signal transduction and adaptation functions of the aspartate receptor in bacterial sensing. *Science* 220:1016-1020.
46. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.
47. Shine, J., and L. Dalgarno. 1975. Determinant of cistron specificity in bacterial ribosomes. *Nature (London)* 254:34-38.
48. Silhavy, T. J., M. L. Berman, and L. W. Enquist. 1984. *Experiments with gene fusions*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
49. Silhavy, T. J., M. J. Casadaban, H. A. Shuman, and J. R. Beckwith. 1976. Conversion of β -galactosidase to a membrane-bound state by gene fusion. *Proc. Natl. Acad. Sci. USA* 73:3423-3427.
50. Tommassen, J., J. Leunissen, M. van Damme-Jongsten, and P. Overduin. 1985. Failure of *E. coli* K-12 to transport *PhoE-LacZ* hybrid proteins out of the cytoplasm. *EMBO J.* 4:1041-1047.
51. Varenne, S., and C. Lazdunski. 1986. Effect of distribution of unfavourable codons on the maximum rate of gene expression by an heterologous organism. *J. Theor. Biol.* 120:99-110.
52. Vieira, J., and J. Messing. 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with universal primers. *Gene* 19:259-268.

VOLUME 170 • JANUARY 1988 • NUMBER 1

Journal of Bacteriology



Published monthly by the
American Society for Microbiology

EN: JOBAAY

170

(1)

1-488

(1988)

ISSN: 0021-9193